

## 統計教育に関する研究動向調査と大学における今後の研究の方向性に関する検討

大和幹枝<sup>1\*</sup>, 松野純男<sup>2</sup>, 土肥弘久<sup>1</sup>, 長南謙一<sup>1</sup>

### A Survey of Research Trends in Statistics Education and Discussion of Future Research Directions at Universities

Mikie Yamato<sup>1\*</sup>, Sumio Matzno<sup>2</sup>, Hirohisa Doi<sup>1</sup>, Ken-ichi Chonan<sup>1</sup>

Statistics education (tokeikyoiku or tokeigakukyoiku in Japanese) is an important part of the educational curriculum, but the topic of statistics can be challenging for some people to learn. This study investigated research trends in statistics education to discuss directions for future research at universities. We analyzed the titles and subtitles of papers published through 2020. The article search was conducted utilizing the Cinii Articles and Japan Medical Abstracts Society databases, and the following search terms were used: "tokeikyoiku" and "tokeigakukyoiku." A total of 1108 papers were analyzed. We visualized trends in the number of papers, and subsequently analyzed the papers using text mining. Results of this analysis suggest that in recent years research efforts have focused more on data utilization, likely due to curriculum guidelines. The results of our analysis highlight the following characteristics related to statistics education at universities: university-related research is not closely related to data utilization research; research is mainly focused on the current situation and not on practical research; and there is a bias between faculties. Future research that corrects these problems is needed to maintain and improve the skills that can be acquired through statistics education.

**Key words:** statistics education, research trends, text mining

Received April 26, 2022; Accepted May 30, 2022

---

<sup>1</sup> Mikie Yamato, Hirohisa Doi, Ken-ichi Chonan 昭和薬科大学臨床薬学教育研究センター

<sup>2</sup> Sumio Matzno 近畿大学薬学部教育専門部門

\* 連絡先：昭和薬科大学臨床薬学教育研究センター 大和幹枝

〒194-8543 東京都町田市東玉川学園 3-3165

Tel: 042-721-1511 Fax: 042-721-1588 E-mail: yamato@ac.shoyaku.ac.jp

## 1. 緒言

統計学は実社会の経済学分野や理工学分野をはじめ様々な分野で応用されている学問である。近年では、ビッグデータと AI (人工知能) を基盤とした問題解決の戦略に統計学は不可欠であり、統計的知見や考え方を教える統計教育<sup>1)</sup>の重要性が非常に高くなっている。実際、初等中等教育における統計教育の充実を目的とし、文部科学省は 2017 年および 2018 年公示の学習指導要領で統計学に関する内容を明示した<sup>2)3)</sup>。このように、国家としてデータサイエンスの進展やデータに基づいた意思決定を重要視していることから<sup>4)</sup>、統計教育は小学校から注力されるようになってきた。

元来、統計教育は、文科系・理科系問わず大学において継続的に取り組まれており、医学部、薬学部および看護学部などの医療系分野においても例外ではない。医療分野では臨床試験に関する論文において統計学が多用されるなど、各分野における統計教育の到達目標は異なるものの<sup>5)</sup>、培われる能力は大学教育において欠かすことは出来ない。しかしながら、統計学の学習は難しく<sup>6)</sup>、苦手な人が少なくないこと<sup>7)</sup>などの理由から統計教育には課題点の散見が予想される。統計学の社会的ニーズを考慮すると、この状況を改善するための研究やその方向性を定める研究が不可欠と思われる。そのためには、まず統計教育の研究動向を明らかにする必要がある。

統計教育の研究動向に関する先行研究では、調査対象となる論文の抽出時に専門誌<sup>8)</sup>やデータベース<sup>9)</sup>が使用されていた。網羅的に論文を抽出するためには、データベースの使用が望ましいと思われるが、データベースを使用した調査研究において、大学での統計教育に関連す

る論文は調査対象外であった<sup>9)</sup>。そこで今回、大学での統計教育に関する今後の研究の方向性を検討するために、初等中等教育の統計教育に関する先行研究を基に、大学における統計教育に関連した論文の調査を試みることにした。

一般的に論文自体はテキストデータに分類されることから、データに含まれる情報を最大限活用するため、研究動向の明確化にあたっては、その解析手法にテキストマイニングを用いた。本手法は、テキストデータに含まれる情報を削ぎ落とすことなく、データ全体を活用できる特徴を有するとされている<sup>10)</sup>。さらに今回の解析では、テキストデータに対する定量的解析とその解析結果を多角的に可視化するために、クラスター分析とヒートマップを組み合わせた解析も行うことにした。

本研究では、これらの手法を用いて統計教育の研究動向を可視化することにより、大学における統計教育に関する今後の研究の方向性について検討を行った。

## 2. 方法

### 1. 解析対象

今回、調査解析の対象は論文表題および副題とした。これは、論文表題が最も端的に論文内容を顕しているため<sup>11)</sup>である。論文の抽出では、伊藤ら<sup>12)</sup>による先行研究と同様に Cinii Articles<sup>13)</sup>および医学中央雑誌 Web 版<sup>14)</sup>(以下、医中誌 Web) の 2 つのデータベースを使用した。

検索語には、「統計教育」および「統計学教育」を使用した(検索式=統計教育 OR 統計学教育)。これは、統計教育に関して、「統計学教育」の用語を用いる文献も見受けられるためである。検索にあたり、「統計教育」および「統計学教育」のシソーラス語、医中誌フリーキー

ワード, 検索支援語<sup>15)</sup>がないことを確認した。

検索期間は, 2020 年までに発表された全ての論文とした(検索年月日:2021年6月28日). Cinii Articles と医中誌 Web には, 原著論文以外にも学会発表などの資料が含まれる<sup>13,16)</sup>. 今回は, 研究動向の主題をより正確に把握するために, 李<sup>17)</sup>と同様に論文の種類は特に制限を設けなかった。

論文表題には, 学会発表での番号や記号, 特集記事やシンポジウム名の情報を含むことが指摘されているため, 畑野<sup>18)</sup>の手法を取り入れ, 論文表題に関係しない情報を削除した。

これらの手順より, Cinii Articles から 1151 件, 医中誌 Web から 42 件の論文を抽出した. その後, 重複していたものを除き, 最終的に 1108 件の論文を解析対象とした(図 1)。

## 2. 解析方法

### 2.1. 論文数の年別推移

論文数の年別推移を可視化(グラフ化)するために, Microsoft Excel 2016<sup>®</sup>を使用した. この際, 解析対象である 1108 件を「統計教育全般」とし, その中に「大学」という語を含む論文数(122 件)を別途抽出, その推移を併記した。

### 2.2. テキストマイニングによるデータ処理方法

テキストマイニングでは, フリーソフトの KH Coder3 b03c<sup>19)</sup>を使用し, 形態素解析には Mecab<sup>20)</sup>を用いた. この理由として, Mecab の利点に他の形態素解析エンジンよりも性能が向上しており, 高速に動作することが挙げられる<sup>21)</sup>. また, 複合語の検出では, 専門用語(キーワード)自動抽出システムである TermExtract<sup>22,23)</sup>を使用し, データの解析にはバックエンドとして R (ver.4.0.5)<sup>24)</sup>を用いた. 尚, 「統計教育」および「統計学教育」の語を

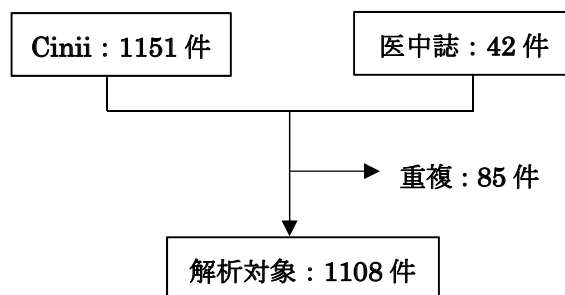


図 1. 論文抽出の過程

検索の段階で使用しており, 解析結果をより精査するために「統計」と「教育」の語は除外した。

### 2.3. 抽出語の頻度集計

形態素解析で抽出された語の頻度集計には Microsoft Excel 2016<sup>®</sup>を使用し, 出現頻度が高い上位 20 語を割り出した. また, 李ら<sup>25)</sup>と同じく, 情報保護の観点から人名および組織名は除外した。

### 2.4. 媒介中心性

統計教育で, 中心的な役割を果たし<sup>26)</sup>, 全体影響が大きい抽出語の検出を行うために媒介中心性<sup>27)</sup>を用いた解析を行った. 論文表題のデータを 1 つのファイルにまとめ, 出現数による取捨選択では最小出現数を 20, 語と語の関連をみるために描画する共起関係の選択では Jaccard 係数を用い<sup>28)</sup>, 描画数は上位 60 に設定した。

### 2.5. 多次元尺度構成法

今回, 研究の特徴を分類する目的で, 多次元尺度構成法を用いた. この手法は, 抽出語の出現頻度をもとにして抽出語同士の関係の把握<sup>29)</sup>を行うためである. 出現数による取捨選択において最小出現数を 20, その他に Kruskal 法,

Jaccard 係数を用い、クラスター数は 4 つとした。尚、クラスター数の設定は、抽出語のクラスター分析の結果を基にして分析者間で決定した。

## 2.6. 対応分析

研究動向の経時的変化の有無を可視化するために、対応分析を行った。対応分析とは、頻度のデータを対象とし<sup>30)</sup>、抽出語の関係性を直感的に把握できる手法である<sup>31)</sup>。今回、年代毎の特徴を概観するために、八軒ら<sup>32)</sup>と同様に 5 年毎に分けて解析した。その後、出現数による語の取捨選択では最小出現数を 20 に設定し、上位 60 語の差異が顕著な語を解析に使用した。

## 2.7. クラスター分析およびヒートマップ

論文発表があった年と抽出語の関連性を可視化するために、クラスター分析およびヒートマップによる解析を行った。クラスター分析は出現パターンの似通った抽出語の組み合わせを探索できる分析であり<sup>33)</sup>、ヒートマップとは抽出語の出現頻度が高くなる程に色が濃くなる手法である<sup>34)</sup>。

縦軸は年、横軸は抽出語とした。まず、年と抽出語の各々で、距離を可視化するためにクラ

スター分析を行った。今回、階層的クラスター分析を使用し、距離の計算は Canberra 距離、クラスタリングでは Ward 法を取り入れた。その後、ヒートマップを作成し、年別の抽出語の出現頻度を可視化した。

## 3. 倫理的配慮

本研究では人を対象としていないため、倫理審査の対象に該当しない<sup>25)</sup>。しかし、著作権の侵害に注意する必要があるため、この対策として李ら<sup>25)</sup>と同様にデータに含まれる個別情報を保護すると共に、解析に使用しないデータは削除した。

## 3. 結果

### 1. 年別の論文数

統計教育全般（解析対象全 1108 件）の論文数の年別推移を図 2 に示した。論文の初出は、1949 年であり、その数は 2000 年頃から急激な増加傾向がみられ、50 件を上回る年が複数あった（2007 年、2009 年、2013 年、2014 年、2018 年）。論文表題に「大学」を含む論文数は、122 件であり、2010 年以降は継続的に報告されていたが、年間 10 件以上の報告された年はなかった（図 2）。

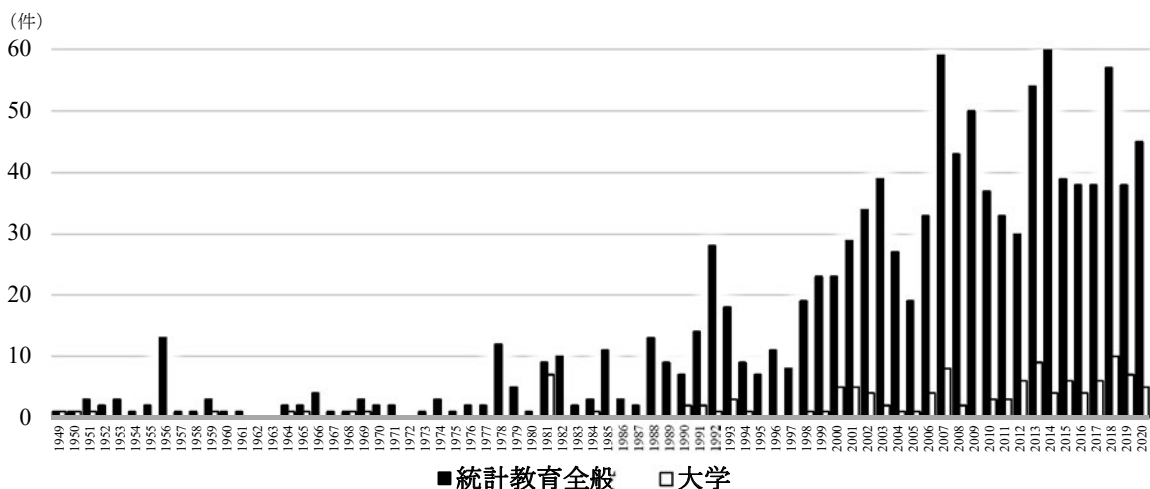


図 2. 論文数の年別推移（統計教育全般および大学）

## 2. 抽出語の頻度集計

抽出語の頻度集計を行った結果、1949年～2020年の計72年分の総抽出語数は13,870語で使用された語数は6,882語、異なり語数は1,714語で使用された語数は1,397語であった。その内、出現頻度が高かった上位20語を表1に示した。その結果、「研究」、「データ」、および「活用」が100回を超える頻出語であり、「大学」の出現順位は上位10位以内であった。一方、医療関連の抽出語としては、「看護」や「医学部」などが出現していたが、上位20位以内には出現していなかった。

表1. 抽出語リスト 上位20語

抽出語	出現回数	抽出語	出現回数
1 研究	130	11 指導	70
2 データ	104	12 教材	68
3 活用	102	13 情報	66
4 実践	93	14 課題	63
5 数学	90	15 カリキュラム	60
6 授業	86	16 分析	57
7 大学	79	17 学生	53
7 問題	79	18 育成	51
9 開発	78	18 学校	51
9 学習	78	18 調査	51

## 3. 媒介中心性

媒介中心性による分析の結果では(図3)、「データ」、「活用」、「授業」、「実践」の中心性が特に高い結果であったが、「大学」は中心性が高い抽出語として検出されなかった。しかし、「大学」は「現状」、「調査」と繋がっており、「大学」と関係が深い「学部」は「文系」と結び付いていた。医療関連の抽出語としては「看護」が表出し、「大学院」や「方法」と繋がっていた。

## 4. 多次元尺度構成法

上記の抽出語同士の関係性を把握するため、多次元尺度構成法の解析では(図4)、媒介中心性が高かった「データ」、「活用」、「授業」、「実践」と「大学」は別のグループであった。「大学」と同一グループであったのは「課題」、「調査」、「現状」であり、近接もしていた。一方、医療関連の抽出語として「看護」が表出し、媒介中心性が高かった抽出語のグループに含まれていたものの、それらの抽出語とは距離が離れていた。

## 5. 対応分析

図5に対応分析を用いて解析を行った結果を示した。経時的に抽出語は変化しており、その年代は大きく3つグループに分かれた。1つ目は、1940年代後半、1950～1960年代、1980年代および2000年代前半であった。この年代付近には、「小学校」、「確率」、「高校」、「大学」、「現状」等が出現していた。2つ目の1970年代および1990年代では「システム」、「ソフトウェア」、「ソフト」、「環境」、「試み」等、3つ目の2000年代後半以降では媒介中心性が高かった「データ」、「活用」、「授業」、「実践」が表出していた。また、抽出語が変化していた時期は、1970年代前半、1990年代前半および2000年代後半という結果が得られた。尚、今回のデータの性質上、1940年代後半は1949年のみ、2020年代前半は2020年のみとなった。

## 6. クラスタ分析およびヒートマップ

上記結果より、文献数が2000年頃から急激に上昇し、対応分析では特に2000年代後半以降で語が頻出していたことから、この年代の傾向をより詳細に把握するために、2000年～2020年の期間の研究動向をクラスタ分析およびヒートマップを用いて可視化した(図6)。



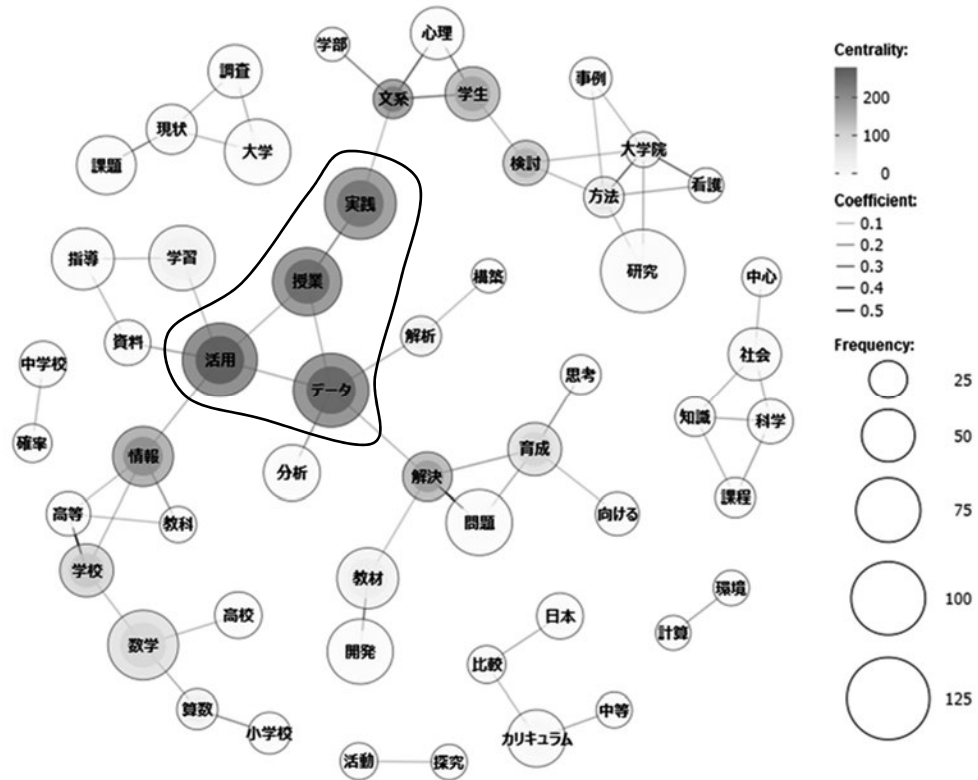


図 3. 媒介中心性

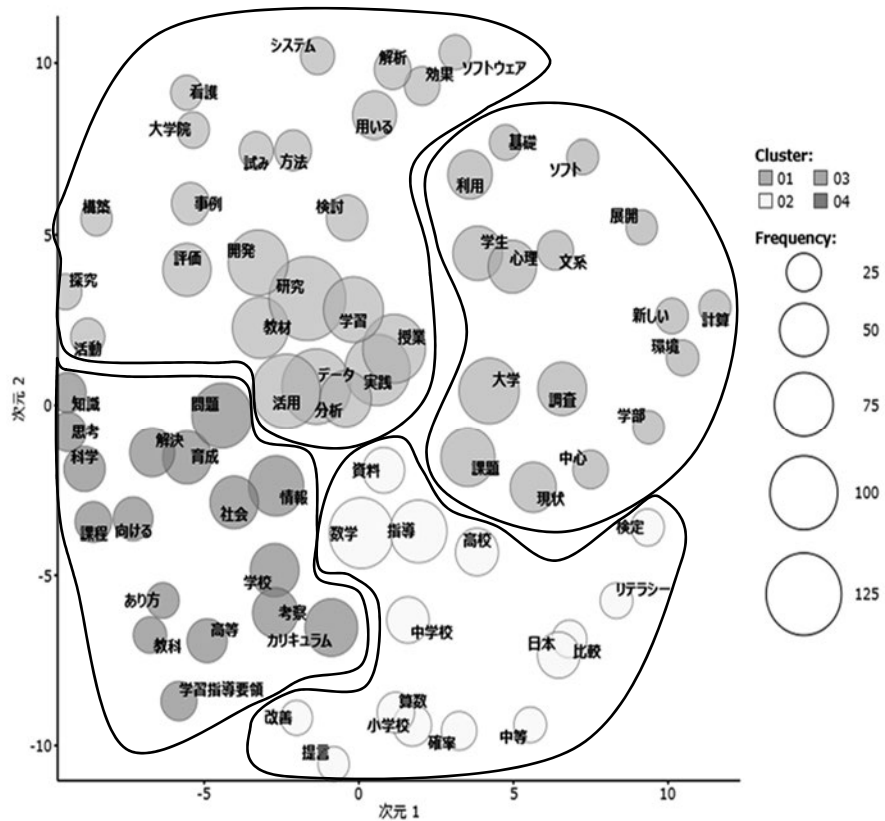


図 4. 多次元尺度構成法

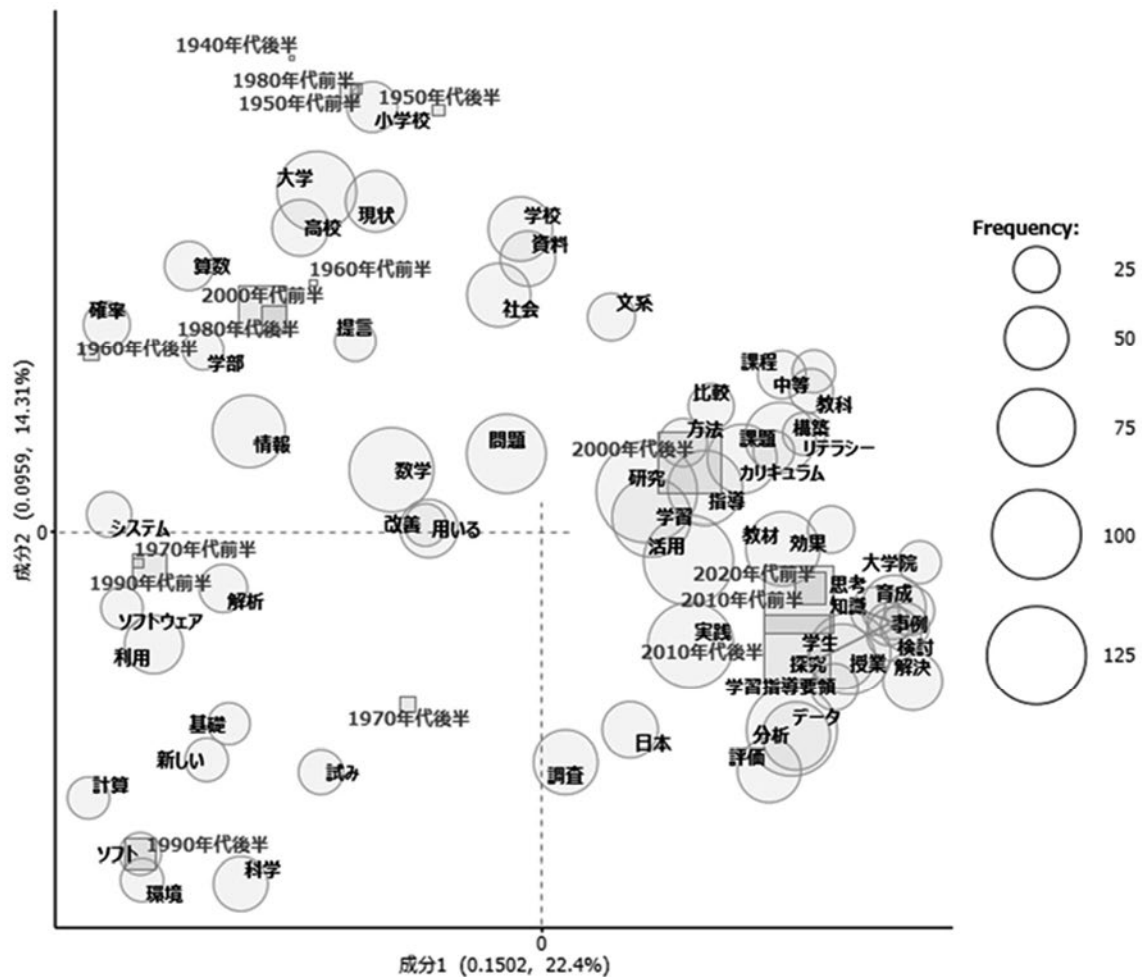


図 5. 対応分析

まず、クラスター分析の結果、期間は大きく2つの群 (A と B) を形成し、A は2000年～2005年、B は2006年～2020年の群となった。

次に、抽出語の分類では、関連性毎に大きく2つに分かれる結果であり (大分類①および②)、大分類①は媒介中心性が高かった「データ」、「活用」、「授業」、「実践」を含んでいた。大分類②では、さらに細分化し、3つのグループを形成していた (小分類②-α、②-βおよび②-γ)。その中に「大学」は小分類②-αに含まれ、医療関連の抽出語としては「看護」が小分類②-βに出現した。したがって、これらの抽出語は、大分類①とは別のグループであることが分かった。

これらの結果をヒートマップにより可視化

すると、抽出語の出現頻度は B 群の年 (2006年～2020年) で高いことが判明した。また、抽出語の分類に着目した場合、大分類①で最も頻出していた。これらより、B 群かつ大分類①のグループで最も抽出語が頻出していることが分かった。

#### 4. 考 察

統計教育に関する論文をデータベースから網羅的に抽出し、研究動向の可視化をテキストマイニングで行った。本研究の全体的な結果として、大学に関する研究の特徴は、大学関連の研究とデータ活用に関する研究との関連性が低いこと、現状調査が中心で実践的な取り組み





が特徴的ではないことおよび検討されている学部への偏りが挙げられる。現在、データ活用に関する取り組みは、初等中等教育で行われており、今後、初等中等教育で得られた素養の維持および向上のために、これらを是正する研究を行うことが望ましいと思われる。以下、結果の詳細について考察する。

まず、年別の論文数（統計教育全般）の結果から、統計教育に関する論文数は徐々に増加しており、統計教育への関心の高さがうかがえる。特に論文数が増加していた2000年頃は、学習指導要領の改訂時期（1998年および1999年）<sup>35)</sup>と重なっていた。また、この改訂では中学校の統計に関する内容が一旦全て削除となったことから<sup>36, 37)</sup>、論文数の顕著な増加傾向に影響した可能性がある。一方、「大学」に関する論文数に急激な増加はなく、検討数の増加が望まれる。

次に、対応分析より、抽出語が年代毎で変わっていたことから、研究動向は年代によって変遷していることが分かった。研究の流れとしては、各教育機関の現状および確率に関する研究またはシステム等に関する研究が交互に繰り返された後、2000年代後半からは教育の具体的な内容に関する研究に集約していると思われる。また、抽出語の特徴が変化していた時期は、1970年代前半、1990年前半および2000年代後半であった。これらのほぼ同時期に学習指導要領の改訂（1968年～1970年、1989年、2008年および2009年）<sup>35)</sup>が行われていたことから、研究動向の変化も学習指導要領の改訂に影響を受けている可能性がある。

2017年の先行研究では、論文抽出の際に専門誌を使用していた検討において、国内の研究は指導や教材およびテクノロジー（電卓、表計算ソフト、統計フリーソフトなど）に大別できることが報告されている<sup>8)</sup>。今回の分析で、指

導や教材は2000年代後半に、テクノロジー関連は1970年代および1990年代で特徴的なことが分かった。

テキストマイニングで得られた抽出語の特徴を捉えるために、頻出語解析、媒介中心性、多次元尺度構成法、クラスター分析の結果を複合的に解析した。それを基に、ヒートマップによる可視化を行った。これらの結果より、統計教育全般で特に頻出し、全体影響が大きいものは、「データ」、「活用」に関する研究と思われる。これは、直近の学習指導要領が影響している可能性があり、実際に小学校で「データの活用」という新たな領域が設けられ、中学校および高等学校においてもこの領域に関する学習の繋がりの考慮が盛り込まれている<sup>2, 3)</sup>。大学に関する研究では、前述した「データ」、「活用」との関連性は低く、主な研究は現状調査であるため、実践的な研究は特徴的なものとして検出されなかったと思われる。加えて、研究が行われている学部への偏りがあったことも、その要因の1つだと思われる。医療関連の研究では、「看護」が突出しており、大学院や方法論に関する研究が中心であることが示唆された。また、「データ」、「活用」との関連性も高いとは言いが難かった。これらの理由として、大学において各学部で求められている能力の違い<sup>5)</sup>が影響している可能性がある。

しかし、ビッグデータから価値を見出すデータサイエンス<sup>38)</sup>が進展する中、どの分野においてもデータを活用する能力は欠かせない。今後、大学関連の研究では、データ活用に関する研究、現状調査で得られた知見を基にした実践的な研究の実施および学部間の偏りを是正することで、統計教育で得られる能力のさらなる伸長を図ることが望ましい。

## 5. 研究の限界

本研究では、海外のデータベースを使用していないため、解析対象の論文は日本語のみであり、論文表題のみを対象としているため、本文は含まれていない。そして、今回の検索語で抽出されない論文は解析対象から外れている。これらは、Lin ら<sup>39)</sup>の報告と同様に選択バイアスに繋がる可能性がある。

## 6. 結 語

本研究結果より、大学におけるデータ活用に関する研究および実践的な研究を実施すること、さらに学部間の偏りを是正する研究を行うことで、学習者が統計教育で修得可能な能力の維持および向上に繋がると思われる。

今後は、国外のデータベースも使用することで、国内外の比較を行う。また、本文も対象に入れることで、各国でどの様な研究が行われているのかを可視化する。それらにより、国内における研究動向のさらなる明確化および今後の研究の方向性に関するより詳細な検討を行う。

## 利益相反

開示すべき利益相反はない。

## 引用文献

- 1) 教育新聞: 【探究的な問題解決能力 (1)】なぜ今統計教育か,  
<https://www.kyobun.co.jp/education-practice/p20191101/>, 2022年4月2日アクセス.
- 2) 文部科学省 初等中等教育局教育課程課: 平成 29 年改訂の小・中学校学習指導要領に関する Q&A<算数,数学に関すること>, [https://www.mext.go.jp/content/1422350\\_001.pdf](https://www.mext.go.jp/content/1422350_001.pdf), 2022年2月25日アクセス.
- 3) 文部科学省 初等中等教育局教育課程課: 平成 30 年改訂の高等学校学習指導要領に関する Q&A<数学に関すること>, [https://www.mext.go.jp/content/1422374\\_001.pdf](https://www.mext.go.jp/content/1422374_001.pdf), 2022年3月14日アクセス.
- 4) 長尾篤志: 統計調査ニュース No.388 新学習指導要領で強化された統計教育のポイント, <https://www.stat.go.jp/info/t-news/pdf/1903.pdf>, 2022年3月14日アクセス.
- 5) 統計関連学会連合 理事会・統計教育推進委員会, 統計教育大学間連携ネットワーク 質保証委員会: 統計学の各分野における教育課程編成上の参照基準, <http://jfssa.jp/ReferenceStandard2.pdf>, 2021年5月7日アクセス.
- 6) Garfield J, How students learn statistics, *International Statistical Review*, 63, 25-34 (1995).
- 7) 伊川美保, 楠見 孝, 統計リテラシー自己効力感尺度日本語版の作成: —統計教育の効果測定—, *心理学研究*, 91, 133-41 (2020).
- 8) 大谷洋貴, 我が国における統計教育研究の傾向, *数学教育学研究*, 全国数学教育学会誌, 23, 33-44 (2017).
- 9) 浅木森利昭, B211 学校教育における統計教育研究の動向と課題, 学校教育研究文献データベースの活用, *日本科学教育学会年会論文集*, 15, 323-4 (1991).
- 10) 那須川哲哉, 吉田一星, 宅間大介, 鈴木祥子, 村岡雅康, 小比田涼介, 1. 2. 2 分類, “テキストマイニングの基礎技術と応用”, 岩波書店, 東京, 2020, pp. 5-7.
- 11) 佐久嶋研, 佐々木秀直, 田代邦雄, テキストマイニングを用いた学会誌論文タイトルの時系列分析—日本神経学会誌「臨床神

- 経学」の分析—, 医療情報学, 32, 315-21 (2012).
- 12) 伊藤 海, 田口敦子, 松永篤志, 竹田香織, 村山洋史, 大森純子, 「互助」の概念分析, 日本公衆衛生雑誌, 67, 334-43 (2020).
- 13) 国立情報学研究所: CiNii Articles—CiNii Articles について, [https://support.nii.ac.jp/ja/cia/cinii\\_articles](https://support.nii.ac.jp/ja/cia/cinii_articles), 2021年7月29日アクセス.
- 14) 特定非営利活動法人 医学中央雑誌刊行会: 医中誌 Web, 医中誌 Web とは, <https://www.jamas.or.jp/service/ichu/>, 2022年3月19日アクセス.
- 15) 特定非営利活動法人 医学中央雑誌刊行会: 医中誌 DB 情報, 索引情報, キーワード, <https://www.jamas.or.jp/database/keyword.html#thesaurus>, 2021年7月29日アクセス.
- 16) 特定非営利活動法人 医学中央雑誌刊行会: 医中誌 DB 情報, 編集方針, 論文種類の定義, <https://www.jamas.or.jp/database/policy2.html>, 2022年4月1日アクセス.
- 17) 李 慧瑛, 日本のメンタルヘルスの研究動向 各年代における労働政策との関連, 日本職業・災害医学会会誌, 68, 348-55 (2020).
- 18) 畑野裕子, 「国際理解教育」の研究動向に関する—考察: CiNii 掲載論文のタイトルに対するテキストマイニングを用いて, 神戸親和女子大学大学院研究紀要 = Bulletin of Kobe Shinwa Women's University Graduate School of Literature, 16, 37-47 (2020).
- 19) 樋口耕一, テキスト型データの計量的分析, —2 つのアプローチの峻別と統合—, 理論と方法, 19, 101-15 (2004).
- 20) 京都大学情報学研究科—日本電信電話株式会社コミュニケーション科学基礎研究所共同研究ユニットプロジェクト: MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://taku910.github.io/mecab/>, 2022年4月20日アクセス.
- 21) 工藤 拓, 1. 3. 3 Mecab, “形態素解析の理論と実装”, 近代科学社, 東京, 2018, p. 8.
- 22) 中川裕志, 前田 朗, 小島浩之: “専門用語(キーワード)自動抽出システム”のページへようこそ, <http://gensen.dl.itc.u-tokyo.ac.jp/>, 2022年4月20日アクセス.
- 23) 樋口耕一, コマンドの概要, “社会調査のための計量テキスト分析: 内容分析の継承と発展を目指して: KH Coder OFFICIAL BOOK: 第2版”, ナカニシヤ出版, 京都, 2020, p. 161.
- 24) R Core Team (2021): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, <https://www.r-project.org/>, 2021年7月30日アクセス.
- 25) 李 慧瑛, 下高原理恵, 峰 和治, 田松裕一, 緒方重光, 医学系文献データベース情報を使ったテキストマイニングの将来展望, 情報の科学と技術, 70, 10, 515-21 (2020).
- 26) 樋口耕一, 中心性とサブグラフ検出, “社会調査のための計量テキスト分析内容分析の継承と発展を目指して: KH Coder OFFICIAL BOOK: 第2版”, ナカニシヤ出版, 京都, 2020, p. 185.
- 27) 角口勝隆: ビッグデータ分析技術を応用したソフトウェア不具合の分析実施事例, <https://www.ipa.go.jp/files/000057839.pdf>, 2021年2月16日アクセス.
- 28) 樋口耕一, 距離, “社会調査のための計量テキスト分析: 内容分析の継承と発展を目指して: KH Coder OFFICIAL BOOK: 第2

- 版”，ナカニシヤ出版，京都，2020，p. 180.
- 29) 末吉美喜, 7.2 アンケートデータの集計と分析, “テキストマイニング入門 Excel と KH Coder でわかるデータ分析”, オーム社, 東京, 2019, p. 204.
- 30) 石田基弘, 金 明哲, 1.4.3 テキストの特徴分析, “コーパスとテキストマイニング”, 共立出版, 東京, 2012, p. 10.
- 31) 末吉美喜, 5.5 関係性を多次元にマッピング?!, “テキストマイニング入門 Excel と KH Coder でわかるデータ分析”, オーム社, 東京, 2019, p. 125.
- 32) 八軒浩子, 松岡有紗, 村井亜衣, 木下佐昌子, 高田充隆, 医療薬学研究のテキストマイニングによる計量的分析「医療薬学」と「日本病院薬剤師会雑誌」の比較, 医薬品情報学, 13, 152-9 (2012).
- 33) 樋口耕一, コマンドの概要, “社会調査のための計量テキスト分析: 内容分析の継承と発展を目指して: KH Coder OFFICIAL BOOK: 第 2 版”, ナカニシヤ出版, 京都, 2020, p. 181.
- 34) Park H., Park M.S., Capturing the trend of mHealth research using text mining, Mhealth, 5, 48 (2019).
- 35) 文部科学省: 資料 7 新学習指導要領関係資料,  
[https://www.mext.go.jp/b\\_menu/shingi/chukyo/chukyo3/004/siryo/\\_icsFiles/afieldfile/2011/04/14/1303377\\_1\\_1.pdf](https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/004/siryo/_icsFiles/afieldfile/2011/04/14/1303377_1_1.pdf), 2022 年 3 月 31 日アクセス.
- 36) 二宮裕之, 統計教育の歴史・現在・今後の課題, 愛媛大学教育学部紀要 第 1 部 教育科学, 50, 123-30 (2004).
- 37) 日本学術会議, 数理科学委員会, 数学教育分科会: 提言 新学習指導要領下での算数・数学教育の円滑な実施に向けた緊急提言: 統計教育の実効性の向上に焦点を当てて,  
<https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-24-t293-2.pdf>, 2022 年 4 月 2 日アクセス.
- 38) 竹村彰通, 和泉志津恵, 齋藤邦彦, 姫野哲人, 松井秀俊, 伊達平和, データサイエンス教育の滋賀大学モデル, 統計数理, 66, 63-78 (2018).
- 39) Lin H. J., Sheu P. C., Tsai J. J. P., Wang C. C. N., Chou C. Y., Text mining in a literature review of urothelial cancer using topic model, BMC Cancer, 20, 462 (2020).